PNNL-15467

The Second Generation Model:
Comparison of SGM and GTAP Approaches to Data Development

Ronald D. Sands[1]
Sebastian Miller[2]
Man-Keun Kim[1]

October
2005

Prepared for the United States Environmental Protection Agency under
Contracts AGRDW89921669 and AGRDW89921964

Joint Global Change Research Institute, College Park, MD
Pacific Northwest National Laboratory
Operated by Battelle for the US Department of Energy

---

[1] Joint Global Change Research Institute, Pacific Northwest National Laboratory, Battelle, University of Maryland, 8400 Baltimore Avenue, Suite 201, College Park, MD 20740, USA

[2] Department of Economics, University of Maryland at College Park

LEGAL NOTICE

This report was prepared by Battelle Memorial Institute (Batelle) as an account of sponsored research activities. Neither Client nor Battelle nor any person acting on behalf of either:

MAKES ANY WARRANTY OR REPRESENTATION, EXPRESS OR IMPLIED, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, product, or composition disclosed in this report may not infringe privately owned rights; or

Assumes any liabilities with respect to the use of, or for damages resulting from the use of, any information, apparatus, process, or composition disclosed in this report.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by Battelle. The views and opinions of authors expressed herein do not necessarily state or reflect those of Battelle.

The Second Generation Model:
Comparison of SGM and GTAP Approaches to Data Development

Introduction

There is a general consensus among economic modelers that information from economic input-output tables should be combined with data on energy quantities for analysis of energy and climate policy. However, at least three different approaches are in use: procedures developed by the Global Trade Analysis Project (GTAP); procedures developed for the Second Generation Model (SGM); and the GTAP-EG approach documented by Rutherford and Paltsev (2000). Here we provide a limited comparison of the GTAP and SGM approaches.[3]

The objectives of this document are to provide a numerical example of the SGM procedure for combining energy balances with input-output data, and compare that example with data from the GTAP data base. We selected China and a base-year of 1997 for this comparison. This year is convenient because it is the base year for GTAP version 5.4, and an input-output table is available for China (with 124 producing sectors) in that year. The GTAP data for China are based on the 1997 China input-output table and on energy balances published by the International Energy Agency (IEA). Therefore, we can demonstrate the SGM procedure using the same input-output tables and energy balances that GTAP begins with.[4]

This document is one of several documents on the Second Generation Model made available during October 2005. We presume that the reader of this document is familiar with both Fawcett and Sands (2005) and Sands and Fawcett (2005), which describe the SGM modeling framework and the procedures used to develop data to populate the SGM model. This document was written in response to a specific query by the Science Advisory Board of the U.S. Environmental Protection Agency to compare SGM data and procedures with those used by GTAP.

We first provide some background on the GTAP data set and procedures. Then we provide a general description of the 1997 input-output table and energy balances for China. The sections after that provide summary statistics from the GTAP data set for China and from the SGM procedure applied to China's input-output table and energy balances. Although the GTAP economic data are all presented in 1997 US$, we can use the 1997 market exchange rate of 8.29 yuan per dollar to compare data sets.[5]

---

[3] We have not included GTAP-EG in this numerical comparison due to time constraints.

[4] The SGM procedure demonstrates how we would proceed if we were to construct a benchmark data set for China with a base year of 1997. This year is also convenient because we avoid the complexity of moving an input-output table across years. Benchmark input-output tables for China are available for 1987, 1992, and 1997.

[5] This is the market exchange rate used by GTAP to convert the 1997 China input-output table from yuan to US dollars (Wang et al., 2002).

Background

The Global Trade Analysis Project (GTAP) began in 1992, and its primary goal is to develop tools and information for global trade analysis. There are two main products from this project. The first one is a global computable general equilibrium model that is built mainly to analyze trade policies, and runs in the GEMPACK modeling environment. The second and most recognized product is the global data set: the most recent version (GTAP 6) has a base year of 2001 with 87 regions and 57 production sectors. The previous version (GTAP 5.4) has a base year of 1997 with 78 regions and 57 production sectors.

The input-output tables provided by GTAP are constructed from three primary data sources: original input-output tables in values, energy quantities from the International Energy Agency, and energy prices from various sources. Therefore, GTAP input-output tables are a hybrid of these sources. There are six energy sectors in the GTAP database: coal, crude oil, natural gas, petroleum and coal products, electricity, and gas manufacture and distribution.

The GTAP procedure produces a balanced set of economic accounts that attempts to reconcile the price, energy quantity, and input-output values: all three data sources are changed in the process of calibration. Data for all GTAP regions are expressed in US dollars, converted from local currencies at market exchange rates.

The SGM approach starts with input-output tables and IEA energy balances, but does not bring independent price information into the calibration process. Energy prices, at least for primary fuels, are determined mainly by production values from the input-output table combined with production quantities from the energy balances. All energy quantities are preserved, but some adjustment to the input-output values is still required. Preservation of energy quantities is essential for analysis of climate policy, as carbon dioxide emissions are linked directly to energy combustion by the ratio of carbon content to energy content in each fuel.

The GTAP-EG approach (Rutherford and Paltsev, 2000) preserves IEA energy quantities and most of the prices, while adjusting the input-output values. Therefore, the SGM and GTAP-EG approaches preserve energy quantities, while the GTAP procedure does not. GTAP-EG is configured for the GAMS modeling environment.

Characteristics of Original Data

Original data sets for this comparison are the 1997 China input-output table and energy balances published by the International Energy Agency. The China input-output table has 124 production sectors, eight of which are energy sectors. See Wang et al. (2002) for details on how the 1997 China input-output table was configured for GTAP. Chinese yuan were converted to U.S. dollars at the 1997 market exchange rate of 8.29 yuan per dollar for inclusion in the GTAP 5.4 data set. For this comparison, two aggregations are used as shown in Table 1, one with 18 production sectors and one with 16.

Table 1. Sector aggregation for 1997 China input-output table. Energy sectors are in bold.

	labels	
Production Sector	China 18	China 16
primary agriculture	PAG	PAG
coal mining and processing	**COL**	**COL**
crude oil	**OIL**	**OIL**
natural gas	**GAS**	**GAS**
food processing	FPR	FPR
wood, paper, pulp	WPP	WPP
petroleum refining	**PRF**	**P_C**
coking	**COK**	
chemical, rubber, plastic products	CRP	CRP
non-metallic minerals	NMM	NMM
iron and steel	I_S	I_S
non-ferrous metals	NFM	NFM
other industry	OID	OID
electricity	**ELY**	**ELY**
steam and hot water	**SHW**	**GTD**
gas production and supply	**GPS**	
transportation	TPT	TPT
services	SVS	SVS

The 18-sector aggregation was selected to match that of a typical SGM region, but also to preserve all energy sectors available from the original input-output table. The 16-sector aggregation combines some of the energy sectors, and corresponds to the way that energy sectors are aggregated in the GTAP input-output table for China. Therefore, the original 1997 China input-output table has eight energy sectors, which appear as six sectors in the GTAP input-output data. The SGM approach preserves greater energy detail, motivated by the energy-CO_2 emissions relationship.

Matching energy balances to economic sectors is straightforward for some energy products, such as coal, but not so simple for others such as natural gas and heat. Energy flows in an energy balance table are complicated and it is helpful to re-organize the data into energy make and use tables. Table 2 contains an energy make table for China, based on 1997 energy balances from IEA. Energy production in China is clearly dominated by coal. Therefore from the perspective of greenhouse gas emissions, coal is an extremely important sector. Not only does it dominate the Chinese energy system, but it also has the highest carbon to energy ratio of any fuel.

Table 2. 1997 energy make table for China (mtoe)

	Crude Oil	Natural Gas	Coal	Coke	Gases from Coal Transform	Refined Petroleum	Electricity	Heat
Oil and Gas Extraction	160.7	21.4						
Coal Mining			713.6					
Coke Ovens				90.3	11.7			
Blast Furnaces					13.7			
Petroleum Refineries						164.5		
Petrochemical Industry								
Gas Works		3.9		0.7				
Public Electricity Plants							95.8	
Autoproducer Electricity Plants							1.7	
Heat Plants								28.1
CHP Plants								
Production by Fuel	160.7	25.3	713.6	91.0	25.4	164.5	97.6	28.1

Notes: Data are from 1997 IEA energy balances for China. Row labels represent energy production and transformation activities. Column labels are energy carriers. No information was available for combined heat and power (CHP) plants. Units for IEA energy data are million tons of oil equivalent (mtoe). One ton of oil equivalent is defined by IEA as 10^7 kilocalories, which equals 41.868 gigajoules (GJ).

Energy balance tables help to sort out the transformations between coal, coke, and natural gas. Various types of natural gas are represented in the energy balances, depending on how they are produced: from oil and gas extraction activities, from coke ovens and blast furnaces, and from gas works (transformation of other fuels to natural gas). It is not immediately clear how to match these physical flows of natural gas to economic flows in the input-output table.

Table 3 presents some basic information from energy balances and the input-output table for five fuels. Definitions of these fuels are relatively consistent between the energy balance table and input-output table. We are particularly interested in the price of fuels implied by the input-output data and energy balances. The column labeled "unit value" is the value of production divided by the quantity produced, and therefore represents an average cost of production.[6]

[6] We use units of dollars per gigajoule for prices (or unit values) of fuels, as this is familiar to most energy modelers.

Table 3. Quantities and values of energy production by fuel.

fuel	production quantity (mtoe)	(PJ)	production value (million yuan)	(million US$)	unit value ($ per GJ)
coal	713.6	29,876	222,748	26,869	0.90
crude oil	160.7	6,730	152,258	18,367	2.73
petroleum products	164.5	6,886	279,576	33,724	4.90
coke	91.0	3,809	30,243	3,648	0.96
electricity	97.6	4,085	377,363	45,520	11.14

Notes: Source of production quantities is IEA energy balances, converted from mtoe to petajoules (PJ) with a conversion factor of 41.868 PJ per mtoe. Source of production values is original 1997 China input-output table. Currency conversion uses 1997 market exchange rate of 8.29 yuan per US dollar.

GTAP Data for China

GTAP version 5.4 contains data for 78 regions and 57 commodities, with a 1997 base year. The data base contains enough information to extract an economic input-output table for each region, and provides information on energy quantities for six fuels. We used the following steps:

1. Aggregate the GTAP data for China to a set of production sectors similar to that used in SGM (16 sectors).
2. Extract GTAP input-output data for China.
3. Extract GTAP energy quantities for China.
4. Provide summary statistics for the energy sectors.

Extracting a balanced input-output table from the GTAP data set requires several steps: the essential reference is McDonald and Thierfelder (2004). We first extracted all the social accounting matrix (SAM) components that are part of an input-output table, and then re-combined them in the format of a typical input-output table.

Input-output data and energy quantities extracted from the GTAP data set are the result of the GTAP fitting procedure. Original data are not provided in the GTAP data set. Background on the use of energy data in GTAP can be found in Complainville and van der Mensbrugghe (1998) and in Burniaux et al. (2002). The GTAP data set includes energy quantities for six fuels. Figure 1 displays the result when energy values are divided by energy quantities, with units of US$ per gigajoule.

Data plotted in Figure 1 represent average prices paid for various fuels by production sectors and households. The price paid is fairly uniform across production sectors and households, indicating that the GTAP data were constructed to adhere generally to the law of one price. This is an important feature: if all consumers of a fuel face the same price, before indirect taxes or other price margins, then a model based on these data can maintain energy balance at each time step. This assures a one-to-one correspondence between values and quantities for each fuel. If consumers face different prices for the same fuel, then energy balance is lost as sectors grow at different rates over time.

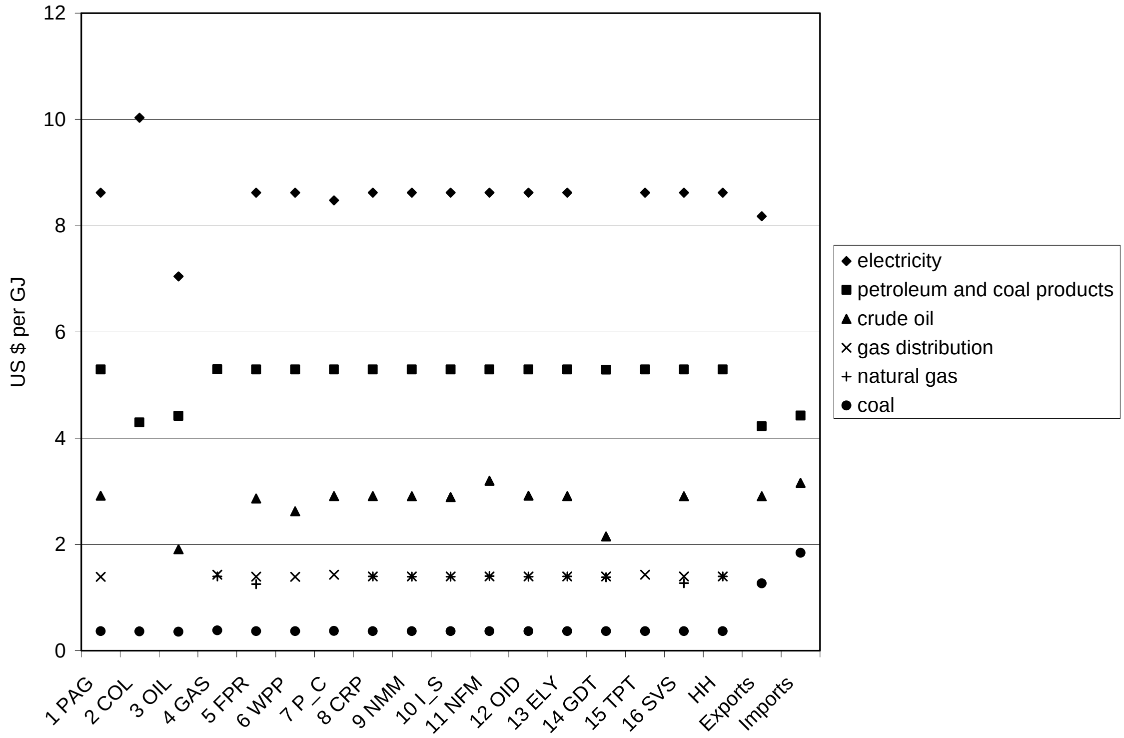

Figure 1. Energy unit values in GTAP data for China. See Table 1 for full names of production sectors.

The numerical value of the price paid for each fuel, for the majority of sectors purchasing that fuel, is shown in the final column of Table 4. The most striking result is that the price paid for coal, at $0.37 per gigajoule (GJ), is much lower in the GTAP data set than the unit value, $0.90 per GJ, from Table 3. This is also reflected in the value of coal production in the GTAP data set, which is much lower than the value of coal production in the original 1997 input-output table for China.

Table 4. Value of production and energy prices (unit values) in GTAP data for China.

	value of production			
fuel	original input-output (million US$)	GTAP 5.4 (million US$)	% change	unit value ($ per GJ)
coal	26,869	11,322	-57.9%	0.37
crude oil	18,367	19,084	3.9%	2.91
natural gas	1,313	752	-42.7%	1.40
gas production and supply	1,683	762	-54.7%	1.40
petroleum and coal products	37,373	37,223	-0.4%	5.29
electricity	45,520	35,213	-22.6%	8.62

Table 5 provides a detailed view of the electricity production sector, both in the original data and from the GTAP data set. The first two columns of numbers in Table 5 are the electricity column from the original China input-output table and from the GTAP input-output table respectively. The final columns contain the energy quantities used in electricity production, both in the GTAP data and in the original IEA energy balances.

Coal provides most of the energy into electricity production, and the quantity of coal used is greater in the GTAP data compared to the original IEA energy balances. Even though the total quantity of coal consumed across all sectors in GTAP is close to the original IEA total, the distribution across sectors is different. Such a large change in the amount of coal used for electricity generation makes it difficult or impossible to link these data to engineering characteristics of coal-fired power plants.

Table 5. Inputs to electricity production from GTAP data.

	original values (million US$)	GTAP values (million US$)	GTAP quantities (mtoe)	IEA quantities (mtoe)
PAG	16	7		
COL	7,940	5,172	336.0	245.0
OIL	882	98	0.8	0.6
GAS	31	144	2.5	3.3
FPR	0	0		
WPP	73	32		
P_C	2,395	4,177	18.9	10.9
CRP	329	168		
NMM	355	170		
I_S	83	41		
NFM	21	11		
OID	6,516	3,326		
ELY	1,525	8,806	24.4	22.7
GDT	30	158	2.7	
TPT	1,347	624		
SVS	4,126	2,021		
primary factors	16,323	7,528		
indirect taxes	3,528	2,729		
TOTAL	45,520	35,213		

It is also interesting to note that a large amount of electricity is consumed in the process of electricity generation, transmission, and distribution. In the IEA energy balances, this is labeled "own use and distribution losses." This shows up in both the original IEA balances and the fitted GTAP data. It appears that the original input-output table does not fully account for own use and distribution losses, and this is an example where the original input-output data and original energy balances do not match.[7] This is corrected in the GTAP data set, with a larger value for electricity input, compared to the original input-output table (in row ELY). However, payments to primary factors are much lower in the GTAP data set than in the original input-output table, and the total value of electricity generated is also lower in the GTAP data.

[7] The value of electricity as an input to electricity production (row ELY) is too low in the original input-output table to account for the large quantity of distribution losses in the IEA energy balances.

SGM Hybrid Procedure

The SGM process for combining an economic input-table with energy balances is described in Sands and Fawcett (2005). The following steps are used in this example:

1. Aggregate the 1997 China input-output table from 124 production sectors to 18 sectors, the same number of sectors as a typical SGM region.
2. Configure the 1997 IEA energy balances for China to match the 18 production sectors in the aggregated China input-output table.
3. Construct a hybrid input-output table for China in 1997.
4. Provide summary statistics for the energy sectors.

For this example, energy balances from IEA are first re-arranged into energy make and use tables, and then combined into a single energy input-output table in a way similar to that of economic make and use tables. See Sands and Fawcett (2005) for background on make and use tables.

The SGM hybrid procedure is designed to enforce the law of one price for each fuel; all producers and consumers face the same price for energy, which is necessary to maintain energy balances as the SGM operates through time. The procedure itself is not very complicated: (1) replace energy rows from the input-output table with energy quantities from the energy balance table; and (2) find a set of implicit prices for each intermediate input to production that rebalances the input-output table. If there are *n* intermediate inputs to production, then we have *n* linear equations in *n* unknowns. The value added component of the input-output table is unchanged.

Table 6 provides summary results from the SGM hybrid table. Of particular interest are the implicit prices (unit values) that are derived from the SGM hybrid procedure. The unit value for coal of $0.87 per GJ is very close to that in Table 3. Unit values for crude oil and petroleum products are also close to the values in Table 3. One exception is the unit value of electricity, which at $14.08 per GJ is somewhat greater than the value of $11.14 in Table 3. This is due almost entirely to own use and distribution losses, which are not fully represented in the original input-output data.

Table 6. Value of production and energy prices (unit values) in hybrid input-output table for China using SGM procedure.

	value of production			
fuel	original input-output (million US$)	hybrid table (million US$)	% change	unit value ($ per GJ)
coal	26,869	25,690	-4.4%	0.87
crude oil	18,367	17,619	-4.1%	2.62
natural gas	1,313	5,548	322.7%	2.61
petroleum products	33,724	34,128	1.2%	4.96
coke	3,648	5,658	55.1%	1.49
electricity	45,520	57,504	26.3%	14.08

Note: Natural gas in the hybrid input-output table is the sum of natural gas and gases from coal transformation.

As was seen in Table 2 there are several ways to produce natural gas, and the complexities of energy accounting for natural gas do not match well with sector definitions in the original input-output table. This is reflected in Table 6 as a large percentage increase in the value of production for natural gas relative to the original input-output table. However, we have a good match in terms of value of production for coal and petroleum products, the two largest fuels in terms of energy consumed in China.

Table 7 provides more detail on the electricity production sector. Note that the SGM hybrid procedure changes none of the energy quantities. Original energy balances are reflected accurately in the final hybrid table. The SGM hybrid procedure has some other nice properties. Value added in the original input-output table is the same as value added in the final hybrid table. Because value added is unchanged for all production sectors, gross domestic product is not affected by the hybrid procedure. Also note that the value of non-energy inputs has not changed much. The increase in the value of electricity production, relative to the original input-output table, is mostly from the large quantity of own use and distribution losses.

Table 7. Inputs to electricity production from SGM hybrid procedure.

	original values (million US$)	hybrid values (million US$)	hybrid quantities (mtoe)	IEA quantities (mtoe)
PAG	16	16		
COL	7,940	8,878	245.0	245.0
OIL	882	70	0.6	0.6
GAS	31	357	3.3	3.3
FPR	0	0		
WPP	73	71		
PRF	2,394	2,268	10.9	10.9
COK	1	0		
CRP	329	338		
NMM	355	332		
I_S	83	84		
NFM	21	22		
OID	6,516	6,365		
ELY	1,525	13,385	22.7	22.7
SHW	27	0		
GPS	4	0		
TPT	1,347	1,429		
SVS	4,126	4,038		
primary factors	16,323	16,323		
indirect taxes	3,528	3,528		
TOTAL	45,520	57,504		

Discussion

The GTAP data set uses a calibration procedure to reconcile three primary sources of data: economic input-output tables, energy quantities, and independent data on prices. The SGM procedure uses economic input-output tables and energy quantities, but does not rely on independent price data. We feel that because the input-output table represents total sales for each fuel, the price derived as the ratio of total sales to total energy in physical units is a better estimate of a price of energy to use across the entire economy.

The SGM approach provides full adherence to energy balances, but with some changes to the economic input-output table. Because the primary focus of the SGM is on analysis of greenhouse gas emissions and the implications of various evolutions of the energy-economy system, we feel that preservation of energy balances is important.

This numerical example focused on fuel prices and the electricity generation sector. Fuel prices matter in the analysis of climate policy, as any carbon price becomes an *additive* price to the price of a fuel, based on the carbon content of the fuel. Therefore, any given carbon price added to a lower (higher) fuel price, becomes a higher (lower) percentage change in the fuel price.

Electricity generation is an important sector in the Second Generation Model because it can be further broken down into various generating technologies and the associated generation and transmission activities. This paper did not go into the details of introducing specific generating technologies to the benchmark data set, but this presents no particular difficulty as long as the collection of individual technologies in the model base year is consistent with base-year energy balances.

The SGM approach allows flexibility in determining how to structure the energy balances so as to minimize loss of information, especially with regard to the number of fuels and energy technologies, and whether to use a source for energy balances other than IEA. The IEA energy balances are not original data, but a compilation of data submitted by individual countries. As with any transformation of data, some information may be lost, a fact that might only come to light when original source data are examined. Our international collaborators are helpful in that regard. With them, we can determine whether to use the IEA energy balances or if there are gains from using locally produced data. Therefore, our international collaborations are useful not only for collecting data, but also for resolving discrepancies.

Whenever possible, SGM data sets remain in local currency and SGM regions operate in local currency. Market exchange rates are applied only as needed for goods traded internationally. We feel that this leaves us in a better position to ultimately address questions related to the difference between market exchange rates and purchasing power parity.

Conclusions

We have compared the GTAP and SGM approaches with a numerical example for China in 1997. The approaches have some similarities. They both use an input-output table and energy quantities to construct a hybrid input-output table. However, the approaches differ as to how the original data sets are reconciled in the final hybrid table, and whether independent data on energy prices are used. This example was designed to provide insights on constructing and updating a benchmark data set for analysis of climate policy. Three of the topics addressed are: the relationship between energy balances and input-output tables; determining the price to use for fuels, especially coal in China; and reasons to maintain strict adherence to energy balances.

IEA energy balances contain a lot of information on the complexity of energy flows and energy transformation, and we have seen that an economic input output table may need adjustment for consistency with energy balances. Both the GTAP and SGM approaches make such adjustments, with distribution losses in the electricity generation sector as an example.

A major difference in results between the two approaches was the final coal price for China. The resulting price from the SGM procedure depends mainly on the economy-wide value of production for coal in the 1997 input-output table for China. When this value is divided by the quantity of coal produced, it should represent an average price of coal for China. GTAP results are influenced by other data on energy prices, and the final coal price is much lower.

The quantity of coal used for electricity generation was 37% greater in the GTAP data set than in the original IEA energy statistics, while the GTAP quantity of total electricity production matches IEA statistics. The GTAP data therefore imply a lower electricity generating efficiency than is present in the original IEA data. The SGM approach of maintaining strict adherence to energy balances reflects the core problem of simulating greenhouse gas emissions over the next several decades. Consistency with energy balances, whether from IEA or another source, is essential for modeling the underlying electricity generating technologies that provide opportunities for reducing greenhouse gas emissions in response to a climate policy.

The GTAP consortium includes a large group of international collaborators that contribute data and provide local insights. For many regions, GTAP will be our only source of input-output information. Finally, we remain open to using any well-constructed data source provided that we obtain a thorough understanding of how the original data were processed and how this process affects our ability to provide analysis relevant to the climate problem

References

Burniaux, J.M., R. McDougall and T.P. Truong. 2002. An Energy Data Base for GTAP, in Dimanaran, B. and R. McDougall eds. *Global Trade, Assistance and Production: The GTAP 5 Data Base*, Center for Global Trade Analysis, Purdue University.

Complainville, C. and D. van der Mensbrugghe. 1998. *Construction of an Energy Database for GTAP V4: Concordance with IEA Energy Statistics*. OECD Development Centre, Paris, France.

Fawcett. A.A. and R.D. Sands. 2005. *The Second Generation Model: Model Description and Theory*. Pacific Northwest National Laboratory, PNNL-15432.

McDonald, S. and K. Thierfleder. 2004. *Deriving a Global Social Accounting Matrix from GTAP Versions 5 and 6 Data*. GTAP Technical Report No. 22, Center for Global Trade Analysis, Purdue University.

Rutherford T. and S. Paltsev. 2000. *GTAP-Energy in GAMS: The Dataset and Static Model*, Center for Economic Analysis Working Paper No. 00-02, Department of Economics, University of Colorado at Boulder.

Sands, R.D. and A.A. Fawcett 2005. *The Second Generation Model: Data, Parameters, and Implementation*. Pacific Northwest National Laboratory, PNNL-15431.

Wang, Z., F. Zhai, and D. Xu. 2002. China (Input-Output Table), in Dimanaran, B. and R. McDougall eds. *Global Trade, Assistance and Production: The GTAP 5 Data Base*, Center for Global Trade Analysis, Purdue University.